Puja Ivaturi

682-699-1060 | puja.datascience@gmail.com | linkedin.com/in/pujaankithaivaturi/ | github.com/PujaAnkithaIvaturi

PROFESSIONAL SUMMARY

A highly analytical Computer Science graduate with a Master's degree, bringing over 3 years of experience in Data Analytics, Machine Learning, and Software Development. Possesses advanced proficiency in programming languages such as Python, SQL, Java, and JavaScript. Enthusiastically seeking opportunities to leverage a diverse skill set in Software Development, CI/CD pipeline creation and maintenance, and Advanced Data Analytics. Demonstrated success in aligning team goals across cross-functional groups, with a focus on uncovering trends, developing innovative products, and providing valuable insights to stakeholders.

EDUCATION

The University of Texas at Arlington Arlington, TX
Master of Science in Computer Science (3.9/4) *Aug.2022 -- Dec.2023*

Ramachandra College of Engineering Eluru, India
Bachelor of Technology in Computer Science (3.7/4) *Jul.2017 – Jul.2021*

EXPERIENCE

Python Developer Feb 2024 – Current
Technoviz LLC *Remote, USA*

- Used Python to create efficient and scalable code, making data processing tasks 30% faster through automation and optimization.
- Utilized important Python libraries like NumPy, Pandas, and Matplotlib to handle data and make it easier to understand.
- Managed MySQL databases, improving data extraction and analysis time by 25%. Designed interactive dashboards using Power BI, providing valuable insights for decision-making.
- Worked on advanced models and collaborated with teams to enhance project adaptability.
- I played a key role in maintaining and optimizing SQL databases, utilizing AWS services for efficient data storage and processing.
- I was responsible for generating detailed financial reports, helping the company make informed decisions about its finances.

Software Engineer, Python Developer Aug 2021 - Oct 2022
Tata Consultancy Services *India*

- Created aggregated daily reports for clients by analyzing credit card transaction data provided by the client and from open data sources, utilizing AWS tools (such as AWS S3 for data storage and AWS Redshift for data warehousing), improving decision-making accuracy by 15%. Use case: Identified fraudulent transactions and unusual spending patterns.
- Utilized statistical methods, metrics, quantitative methods, strategic project plans for data collection, and AWS Glue for data preparation, ensuring data integrity and accuracy. Use case: Ensured accurate customer credit scoring and risk assessment.
- Analyzed credit card data from various clients and sources, visualized findings in dashboards using tools like Tableau and AWS QuickSight, and presented to upper management, resulting in a 25% improvement in decision-making efficiency. Use case: Provided insights on customer spending behavior and preferences.
- Developed and maintained statistical dashboards in Tableau and AWS QuickSight to track bi-weekly sprint task progress based on KPIs, increasing team efficiency by 20%. Use case: Monitored and optimized the performance of credit card processing systems.
- Analyzed and optimized complex SQL queries for various credit card products, reducing total execution time by 13% across all products. Utilized Amazon RDS for database management to enhance performance. Use case: Improved the speed and efficiency of transaction processing and reporting.
- Utilized visualization tools such as Power BI, Tableau, AWS QuickSight, and SQL dashboards for data searching, exploration, and visualization, improving data analysis efficiency by 20%. Use case: Enhanced the ability to track and report on key financial metrics, such as transaction volumes and revenue.

Web Developer Jun. 2018 – Jan. 2020
Ramachandra College of Engineering *India*

- Developed a web application for the college including all the information about the college, events, various departmental data and updates related, and hosted it on Apache Tomcat server and further deployed it on AWS platform.
- Utilizing agile development methodology, I worked with a team of software developers on sprints.
- Developed UML diagrams and design requirements based on provided user needs, document solutions and presented weekly reports.
- Closely dealt with the client's inputs and made sure of delivering the project in time successfully.

PROJECTS

California House Prediction | *Python, Linear Regression, Flask, gunicorn, Colab, Postman, Heroku, Git*
- Developed a California House Pricing Prediction model using Linear Regression, achieving an 59.5% of r_2.
- Implemented Machine Learning Ops, CI/CD pipelines for automation, a Flask web application with a user friendly interface, enabling users to input house features and receive realtime price predictions.
- Deployed the model on Heroku, utilizing Git for version control and Postman for API testing, demonstrating proficiency in end-to-end machine learning application development.

BigMart Sales Prediction | *Matplotlib, Seaborn, Pandas, Numpy, Imputations, XGBoost, RandomForest, Tkinter, Joblib*
- Implemented a sales prediction project using machine learning techniques for Big Mart, showcasing proficiency in data analysis and predictive modeling using Python, Pandas, NumPy, and Seaborn to explore and analyze a dataset containing information about various products and their sales in different outlets.
- Employed data preprocessing techniques, including handling missing values through interpolation and mode-based imputation, feature engineering to extract relevant information, and encoding categorical variables using Ordinal Encoder.
- Utilized XGBoost Random Forest Regressor to predict sales, achieving an R-squared score of approximately 0.60 through cross-validation.

Automatic Viva Conductor | *Python – Tkinter, Excel*
- Engineered this application in Python with Tkinter, reducing manual data entry by 80% and streamlining report generation.
- Achieved a 50% reduction in time spent on viva examination processes, enhancing efficiency and objectivity.

Plant Leaf Disease Detection | *CNN,ML,Activation Functions*
- Detection and Classification of diseases in Plant Leaves. Used Convolution Neural Network for Mathematical Operations with inputs like image shape, width ,height and depth
- Activation Function, MaxPooling, Flatten, Full Connection also used as the functions

Modified version of Nim called red-blue Nim against a human player | *Python, Alpha Beta Pruning*
- Human's move is taken as input.
- Computer's moveis Calculated using Min-Max Algorithm with alpha-beta pruning.

Search Engine | *NLTK, NLP*
- Developed a powerful search engine leveraging cosine similarity scores and NLTK libraries for precise and efficient document retrieval.
- Implemented advanced natural language processing techniques to enhance search accuracy, relevance of results.
- Optimized the search engine's performance by continuously refining algorithms and data preprocessing methods.

Facial Monitoring and 3D Reconstruction | *Open CV, Open3D, Tensorflow, Dlib, CNN, Deep Learning, LLM*
- Developed a Software Program Capable of recognizing facial traits from photos, generate a 3Dmodel of the face.
- Implementation utilizes Python Programming Languages and various Libraries such as OpenCV for image processing
- Dlib for Facial Landmark detection , Tensor Flow, Open 3D for 3D Model Creation.
- CelebA dataset is used to train a Convolutional Neural Network (CNN) for facial trait Classification.

TECHNICAL SKILLS

Languages:Python, R, Java, C/C++, SQL (Postgres, MySQL, MongoDB), JavaScript, HTML/CSS, React Js, Django, Flask
Concentrated Areas: Data Science, Machine Learning, Data Analytics, Full Stack Dev, Cloud (AWS)
Developer and Project Tools: Git, Docker, Google Cloud Platform, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse, Jira
ETL, Reporting Tools: Tableau, PowerBI, Weka, Google Sheets, Excel
Data Processing and Streaming: Pandas, NumPy, Matplotlib, Scikit-learn, Seaborn, Tensorflow, Pyspark, Keras, Pytorch, Pickle, PyTest, Stack: "Apache Kafka, PySpark, Apache Airflow, Glue, Athena, Databricks"
Machine Learning: Linear Regression, Logistic Regression, Decision Tree, SVM, K mean, Random Forest
Cloud: AWS(S3, EC2, Lambda, IAM, RedShift, Athena, Glue), Azure(Data Lake, Data Blob Storage, Databricks), Spark, Kafka
Other Technical Skills: Data Modeling, Data Warehouse, Data Quality, Data Analysis, Data Governance, Team Collaboration

PUBLICATIONS

"Scope of e-learning in India", IJEAT, 2019: A data research and analysis paper over e-learning in India.
"Grays of Life", 2020: Describes a combination of experiences one cherishes throughout their life.
"A project paper: Digital Examination Set-up", IJIEMR, 2021.: A project paper on how we transformed the paper taking exams into digital using web technologies, and how it impacted our examination system
"Stimulus: A Medical Humanities Journal", 2023.
45M reach in Brainly: My educational content on Brainly has so far helped 45M+ students.